

# HunyuanVideo: A Systematic Framework For Large Video Generative Models

Year: 2024 | Citations: 671 | Authors: Weijie Kong, Qi Tian, Zijian Zhang, Rox Min, Zuozhuo Dai

---

## Abstract

Recent advancements in video generation have significantly impacted daily life for both individuals and industries. However, the leading video generation models remain closed-source, resulting in a notable performance gap between industry capabilities and those available to the public. In this report, we introduce HunyuanVideo, an innovative open-source video foundation model that demonstrates performance in video generation comparable to, or even surpassing, that of leading closed-source models. HunyuanVideo encompasses a comprehensive framework that integrates several key elements, including data curation, advanced architectural design, progressive model scaling and training, and an efficient infrastructure tailored for large-scale model training and inference. As a result, we successfully trained a video generative model with over 13 billion parameters, making it the largest among all open-source models. We conducted extensive experiments and implemented a series of targeted designs to ensure high visual quality, motion dynamics, text-video alignment, and advanced filming techniques. According to evaluations by professionals, HunyuanVideo outperforms previous state-of-the-art models, including Runway Gen-3, Luma 1.6, and three top-performing Chinese video generative models. By releasing the code for the foundation model and its applications, we aim to bridge the gap between closed-source and open-source communities. This initiative will empower individuals within the community to experiment with their ideas, fostering a more dynamic and vibrant video generation ecosystem. The code is publicly available at <https://github.com/Tencent/HunyuanVideo>.